

## DOCUMENT RESUME

ED 329 569

TM 016 193

AUTHOR Spray, Judith A.; And Others  
TITLE Comparison of Two Logistic Multidimensional Item Response Theory Models. Research Report ONR90-8.  
INSTITUTION American Coll. Testing Program, Iowa City, Iowa.  
SPONS AGENCY Office of Naval Research, Arlington, VA. Cognitive and Neural Sciences Div.  
REPORT NO ACT-RR-90-8; ONR90-8  
PUB DATE Oct 90  
CONTRACT N00014-89-J-1908  
NOTE 47p.  
AVAILABLE FROM ACT Research Report Series, P.O. Box 168. Iowa City, IA 52243.  
PUB TYPE Reports - Research/Technical (143)  
  
EDRS PRICE MF01/PC02 Plus Postage.  
DESCRIPTORS Ability; Comparative Analysis; \*Item Response Theory; \*Mathematical Models; \*Scores; \*Test Items  
IDENTIFIERS Ability Estimates; \*Multidimensional Models

## ABSTRACT

Test data generated according to two different multidimensional item response theory (IRT) models were compared at both the item response level and the test score level to determine whether measurable differences between the models could be detected when the data sets were constrained to be equivalent in terms of item "p"-values. The models chosen were the multiplicative multidimensional IRT model (non-compensatory) and a linear multidimensional IRT model (compensatory). Data on a 20-item test for 2,000 examinees were used. Although differences could be detected at the item level, these differences decreased as the correlation between examinee abilities increased. Furthermore, these item differences were small and could be considered unimportant or insignificant from a practical standpoint. No differences were found at the total test score level, and it was concluded that, at least for the data used in this study, the models were indistinguishable. Nine tables and eight figures present study data. (Author/SLD)

\*\*\*\*\*  
\* Reproductions supplied by EDRS are the best that can be made \*  
\* from the original document. \*  
\*\*\*\*\*

TM

ED329569

---

## ACT Research Report Series

---

90-8

U.S. DEPARTMENT OF EDUCATION  
Office of Educational Research and Improvement  
EDUCATIONAL RESOURCES INFORMATION  
CENTER (ERIC)

☒ This document has been reproduced as  
received from the person or organization  
originating it.  
☐ Minor changes have been made to improve  
reproduction quality.

• Points of view or opinions stated in this docu-  
ment do not necessarily represent official  
OERI position or policy.

# Comparison of Two Logistic Multidimensional Item Response Theory Models

## Research Report ONR90-8

Judith A. Spray, Tim C. Davey, Mark D. Reckase  
Terry A. Ackerman  
James E. Carlson

Prepared under Contract No. N00014-89-J-1908, Contract Authority  
Identification No. NR 154-531, with the Cognitive Science Research Program of  
the Office of Naval Research.

Approved for public release; distribution unlimited. Reproduction in  
whole or in part is permitted for any purpose of the United States Government.

---

October 1990

**BEST COPY AVAILABLE**

**ACT**

For additional copies write:  
ACT Research Report Series  
P.O. Box 168  
Iowa City, Iowa 52243

©1990 by The American College Testing Program. All rights reserved.

**Comparison of Two Logistic Multidimensional Item  
Response Theory Models**

**Judith A. Spray, Tim C. Davey, and Mark D. Reckase**  
**American College Testing**

**Terry A. Ackerman**  
**University of Illinois**

**James E. Carlson**  
**Auburn University at Montgomery**

Approved for public release; distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the United States Government.

REPORT DOCUMENTATION PAGE				Form Approved OMB No 0704 0188	
1a REPORT SECURITY CLASSIFICATION <b>UNCLASSIFIED</b>			1b RESTRICTIVE MARKINGS		
2a SECURITY CLASSIFICATION AUTHORITY			3 DISTRIBUTION / AVAILABILITY OF REPORT <b>Approved for public release: distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the U.S. Govt.</b>		
2b DECLASSIFICATION / DOWNGRADING SCHEDULE					
4 PERFORMING ORGANIZATION REPORT NUMBER(S) <b>ONR 90-8</b>			5 MONITORING ORGANIZATION REPORT NUMBER(S)		
6a NAME OF PERFORMING ORGANIZATION <b>ACT</b>		6b OFFICE SYMBOL (If applicable)	7a NAME OF MONITORING ORGANIZATION <b>COGNITIVE SCIENCE RESEARCH PROGRAMS OFFICE OF NAVAL RESEARCH</b>		
6c ADDRESS (City, State, and ZIP Code) <b>P.O. Box 168 Iowa City, IA 52243</b>			7b ADDRESS (City, State, and ZIP Code) <b>Code 1142CS Arlington, VA 22217-5000</b>		
8a NAME OF FUNDING / SPONSORING ORGANIZATION		8b OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER <b>N00014-89-J-1908</b>		
8c ADDRESS (City, State, and ZIP Code)			10 SOURCE OF FUNDING NUMBERS		
			PROGRAM ELEMENT NO <b>61153N</b>	PROJECT NO <b>RR04204</b>	TASK NO <b>RR0420401</b>
WORK UNIT ACCESSION NO <b>4421556--- 02</b>					
11 TITLE (Include Security Classification) <b>Comparison of two logistic multidimensional item response theory models</b>					
12 PERSONAL AUTHOR(S) <b>Judith A. Spray, Tim C. Davey, Mark D. Reckase, Terry A. Ackerman, James E. Carlson</b>					
13a TYPE OF REPORT <b>Technical</b>		13b TIME COVERED FROM _____ TO _____		14 DATE OF REPORT (Year, Month, Day) <b>1990, October</b>	
15 PAGE COUNT <b>37</b>					
16 SUPPLEMENTARY NOTATION					
17 COSATI CODES			18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB GROUP	<b>Multidimensional Item Response Theory, Compensatory Models, Noncompensatory Models</b>		
<b>05</b>	<b>09</b>				
19 ABSTRACT (Continue on reverse if necessary and identify by block number) <b>Test data generated according to two different multidimensional item response theory models were compared at both the item response level and the test score level to determine if measurable differences between the models could be detected when the data sets were constrained to be equivalent in terms of item p-values. Although differences could be detected at the item level, these differences decreased as the correlation between examinee abilities increased. Furthermore, these item differences were small in magnitude and could be considered unimportant or insignificant from a practical standpoint. No differences were found at the total test score level, and it was concluded that, at least for the data used in this study, the models were indistinguishable.</b>					
20 DISTRIBUTION / AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT <input type="checkbox"/> DTIC USERS			21 ABSTRACT SECURITY CLASSIFICATION <b>UNCLASSIFIED</b>		
22a NAME OF RESPONSIBLE INDIVIDUAL <b>Dr. Charles Davis</b>			22b TELEPHONE (Include Area Code) <b>(703) 696-4046</b>		22c OFFICE SYMBOL <b>ONR 1142CS</b>

## **Abstract**

Test data generated according to two different multidimensional item response theory models were compared at both the item response level and the test score level to determine if measurable differences between the models could be detected when the data sets were constrained to be equivalent in terms of item  $p$ -values. Although differences could be detected at the item level, these differences decreased as the correlation between examinee abilities increased. Furthermore, these item differences were small in magnitude and could be considered unimportant or insignificant from a practical standpoint. No differences were found at the total test score level, and it was concluded that, at least for the data used in this study, the models were indistinguishable.

## Comparison of Two Logistic Multidimensional Item Response Theory Models

Psychometricians who have some interest in multidimensional item response theory (MIRT) modeling may be familiar with the terms, *compensatory* and *noncompensatory* as they relate to two general model classification schemes. Ansley and Forsyth (1985) contrasted the two types of model classifications as follows. "Compensatory models, unlike noncompensatory models, permit high ability on one dimension to compensate for low ability on another dimension in terms of probability of correct response. In the noncompensatory models, the minimum factor (probability) in the denominator is the upper bound for the probability of a correct response. Thus, for a two-dimensional item, a person with a very low ability on one dimension and very high ability on the other has a very low probability of correctly answering the item" (p. 40).

Typically, MIRT models of the compensatory type, such as the logistic MIRT model (Doody-Bogan & Yen, 1983; Hattie, 1981; Reckase, 1985, 1986) or the normal ogive MIRT model (Samejima, 1974) imply linear combinations of the multidimensional abilities in the exponent of the expression for the probability of a correct response. In this linear fashion, a low ability on one or more of the  $k$  ability dimensions can be compensated by a higher ability on one or more of the remaining dimensions. Because the compensation is a characteristic of this linear combination, such models are probably more accurately labeled *linear* MIRT models. A typical linear logistic MIRT model of the compensatory type can be written as

$$P_j(\theta_i) = c_j + (1 - c_j) \frac{e^{\left\{ \sum_{m=1}^k f_{ijm} \right\} \cdot d_j}}{1 + e^{\left\{ \sum_{m=1}^k f_{ijm} \right\} \cdot d_j}} \quad (1)$$

where

$$f_{ijm} = a_{jm} \theta_{im}$$

$c_j$  = the pseudo-guessing parameter of the  $j$ th item,

$a_{jm}$  = the discrimination parameter for the  $j$ th item on the  $m$ th dimension,

$d_j$  = the difficulty parameter for the  $j$ th item, and

$\theta_{im}$  = the  $m$ th element in the  $i$ th person's ability vector,  $\theta_i$ .

In this model the favorable response probability,  $P_j(\theta_i)$ , is bounded from below by  $c_j$ . However, because the upper bound of  $P_j(\theta_i)$  is not a function of any one ability dimension, it increases monotonically as  $\sum_{m=1}^k f_{ijm}$  increases.

On the other hand, noncompensatory MIRT models (Simpson, 1978; Embretson, 1984) describe the probability of a favorable response in terms of a product of  $k$  functions of ability on a single dimension and item characteristics. In its most common form, a logistic MIRT model of this noncompensatory or multiplicative type can be written as

$$P_j(\theta_i) = c_j + (1 - c_j) \prod_{m=1}^k \frac{e^{f_{ijm}}}{(1 + e^{f_{ijm}})}, \quad (2)$$

where now we let  $f_{ijm} = [a_{jm} (\theta_{im} - b_{jm})]$  with  $b_{jm}$  = the difficulty parameter for the  $j$ th item on the  $m$ th dimension.  $P_j(\theta_i)$  is bounded by an upper asymptote equal to the minimum of  $\exp\{f_{ijm}\}/(1 + \exp\{f_{ijm}\})$ , and the lower asymptote,  $c_j$ , for any given examinee with  $\theta = \theta_i$ . Thus, the noncompensatory nature of the model is due to the fact that  $P_j(\theta_i)$  can never be greater than the minimum value of the terms in the product,  $\exp\{f_{ijm}\}/(1 + \exp\{f_{ijm}\})$ , a function of the smallest value of the  $k$



ability dimensions for a given examinee. Because of its multiplicative form, the model is more generally labeled as a *multiplicative* MIRT model.

Researchers have used the multiplicative MIRT model to examine characteristics of unidimensional item response theory parameter estimates derived from MIRT response data (Ansley & Forsyth, 1985) and to model certain multicomponent latent traits in response processes (Embretson, 1984). Reckase (1985) has used a linear MIRT model on real response data to estimate two-dimensional item and person parameters on an ACT Assessment Mathematics Usage test. However, no one has actually shown that one model is more representative of the actual item-examinee response process than the other. It may even be possible that one model may be appropriate under one set of circumstances while the other type may be more appropriate in other situations.

In this paper we investigate the differences between item responses generated by these two logistic MIRT models. We have been interested in determining whether or not it is possible to distinguish one model or process from the other through some evaluation of response data. More specifically, our concern has been in establishing whether or not it is possible to detect differences between these two MIRT models, either at the item response or test score level, when the item parameters from each MIRT model have been matched or equated in some sense.

The first task was to establish the item parameters from one of the logistic MIRT models that would produce "reasonable"  $p$ -values or proportion-correct indices for a specified examinee population. Therefore, a target distribution of  $p$ -values for a 20-item test was conceived and item parameters for a linear or compensatory MIRT model were chosen, basically by trial-and-error, until the expected  $p$ -value with respect to this examinee population matched the target distribution. Table 1 gives the set of item parameters for the 20 items for the model given by equation (1). The table also gives the expected value of each  $p$ -value under the assumption that the ability vector,  $\theta$ , for the examinee population,

was distributed as bivariate normal with mean vector,  $\mathbf{0}$ , and variance-covariance matrix of ones along the diagonal and with nondiagonal values equal to rho (.00, .25, .50, or .75). All c-parameters were set to zero.

.....  
 Insert Table 1 Here  
 .....

In order to produce a comparable or "matched" set of noncompensatory, or multiplicative model item parameters, estimates of these item parameters were obtained by minimizing

$$\sum_{i=1}^N \{ [ P_C(\theta_i, a, d) ] - [ P_{NC}(\theta_i, \hat{a}, \hat{b}) ] \}^2 \quad (3)$$

for  $N = 2000$  randomly selected examinees with ability,  $\theta$ , distributed as given previously, where  $P_C$  and  $P_{NC}$  represent logistic MIRT models given by equations (1) and (2), respectively. This process was repeated for 10 replications for each of  $k = 1, 2, \dots, 20$  items to insure that the estimates obtained weren't unduly influenced by the samples selected or the starting values used. Mean values of the replication estimates yielded the noncompensatory item parameters listed in Tables 2-5, for rho values of .00, .25, .50, and .75. The expected value of each item's  $p$ -value is given in the last column of each table. Because the least squares minimization procedure produces unbiased estimates of  $P_{NC}$ , the expected value of each  $p$ -value under the noncompensatory model should be equal to that of the compensatory model, within some estimation error. Equivalence of  $p$ -values was the critical matching criterion between the two MIRT models.

.....  
 Insert Tables 2-5  
 Here  
 .....

### Model Differences at the Total Test Level

By treating the two sets of item parameters as known for each of the two MIRT models, we first investigated the differences between expected number-correct score frequencies of a 20-item test when  $\theta$  was distributed as a bivariate normal random vector with distributions given previously. These frequencies were estimated by evaluating either the number-correct distribution under the compensatory model,  $h_C(y)$  or the noncompensatory model,  $h_{NC}(y)$ , for  $y = 0, 1, 2, \dots, 20$ , or

$$h_C(y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_C(y | \theta) g(\theta) d\theta_1 d\theta_2 \quad (4)$$

and

$$h_{NC}(y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{NC}(y | \theta) g(\theta) d\theta_1 d\theta_2 . \quad (5)$$

In each case, the conditional frequencies,  $f_C(y | \theta)$  and  $f_{NC}(y | \theta)$ , were computed using either models (1) or (2), and a recursive procedure described by Lord and Wingersky (1984). Table 6 gives the signed differences between the frequencies,  $h_C(y) - h_{NC}(y)$ , for  $y = 0, 1, 2, \dots, 20$ , for rho values of .00, .25, .50, and .75. The greatest differences, as expected, occurred for the highest number-correct scores, but the differences in frequencies were small, never greater than .015. For most number-correct score values, these differences became smaller as rho increased.

Another way to assess the significance of these differences was to determine how much data would need to be observed before the differences were statistically detectable. This was done by calculating the minimum sample size required to reject the homogeneity of parallel populations with given levels of test significance and power. These calculations assumed a multivariate normal approximation for each model's multinomial distribution of observed-score frequencies which in turn produced the quadratic form of the noncentrality parameter of a noncentral chi square distribution. The minimum sample size followed as a direct function of this parameter, the specified test significance, and power. For example, with a significance level of .01 and power equal to .95, the minimum sample sizes were 1678, 3242, 7466, and 15311 for correlated ability distributions with rho equal to .00, .25, .50, and .75, respectively. These sample sizes state that even in the unlikely event of uncorrelated ability distributions, it would *still* require at least 1678 observed scores from both the compensatory *and* noncompensatory MIRT models before the null hypothesis of model equivalence could be rejected with a power of .95.

.....

Insert Table 6 Here

.....

The first four (central) moments of each number-correct distribution are given in Table 7 for each value of rho. Both distributions were negatively skewed with the compensatory distribution slightly more platykurtic and both were generally flatter than the normal distribution. The variances of the number-correct scores increased with an increase in rho, and in general, the distributions of number-correct scores became increasingly similar as rho increased.

.....

Insert Table 7 Here

.....

A contour plot of the (signed) difference between the number-correct true scores under the two models, or

$$\sum_{j=1}^{20} P_{jC}(\theta) - \sum_{j=1}^{20} P_{jNC}(\theta)$$

was another way to observe model differences at the total test level for various  $(\theta_1, \theta_2)$  points in the ability space. The greatest differences occurred when either  $\theta_1$  or  $\theta_2$  was low. See Figures 1-4 for rho values of .00, .25, .50, and .75, respectively. It should be noted that, in these plots, the only influence of rho was through the values of the noncompensatory item parameters. Recall that the compensatory item parameters were fixed for all values of rho. Therefore, when interpreting these contour plots, one has to mentally superimpose the appropriate bivariate normal distribution over the contours in order to evaluate the importance of the true-score differences observed.

.....

Insert Figures 1-4  
Here

.....

Another way to compare the two MIRT models was to observe the amount of multidimensional information (MINF) for different points in the ability space between the two models. MINF has been defined (Reckase, 1986) as a direct generalization of the unidimensional IRT concept of item information (i.e., the ratio of the square of the slope of the item characteristic curve at an ability point,  $\theta$ , to the variance of the error of the item score at that level of  $\theta$ ). For the definition of MINF, the slope of the item characteristic surface must be evaluated in a particular direction,  $\alpha$ , a vector of angles with the coordinate axes of the ability space.

Plots of the absolute difference between the compensatory and noncompensatory test information vectors (i.e., the sum of item information across

the 20 items) for item parameters estimated with rho values of .00, .25, .50, and .75 (Figures 5-8, respectively) showed that model differences might be significant if abilities were negatively correlated. However, for all "likely" ability distributions, there were no meaningful differences in MINF between the two models, and these absolute differences appeared to decrease as rho increased.

---

Insert Figures 5-8 Here

---

### Model Differences at the Item Level

It was also of interest to evaluate the differences between models at the single item response level. There were two ways in which this was done. The first involved the evaluation of the *ideal observer index* (Davey, Levine, & Williams, 1989; Levine, Drasgow, Williams, McCusker, & Thomasson, 1990). A more complete definition of this index is provided in the appendix of this paper. However, a simplified definition is as follows. The *ideal observer index* (IOI) is a measure of the proportional number of times that a correct decision is made concerning which of the two competing models produced a particular response to an item. The decision is one that is made hypothetically by an "ideal observer," or an individual who has access to all of the information necessary to yield the highest possible percent of model classification (i.e., compensatory vs. noncompensatory). As far as the ideal observer is concerned, if the item response data fail to distinguish between the two competing models, then the value of this index would be at or near the chance level of .5. Conversely, readily distinguishable models should yield an index near 1.0.

Table 8 shows that the IOI was greater than chance, implying that there was a difference between the models for all 20 items. However, the IOI was never greater than .60 and was greater than .55 for only three items, numbers 3, 6,

and 7, when rho was .00. The value of the IOI decreased for each item as rho increased, implying that it became more difficult to distinguish between the models as the correlation coefficient increased.

One way to think of the magnitude of the IOI was to imagine how many trials of the IO experiment would be necessary before the ideal observer could ascertain, with some given level of certainty, that the models were actually distinguishable. This would be comparable to a test of the difference between any obtained IOI from Table 8 and the null proportion of correct model classifications due to chance. For example, to be able to detect a true difference between the models for item number 6 with a zero value of rho would require at least 40 trials of the IO experiment. This would be comparable to a test of the null proportion of correct classifications due to chance or .50 versus the (true) alternative proportion (.555) with a significance of .01 and power of .95. Conversely, a true IOI of .52 would require more than 290 trials at similar levels of test significance and power.

.....

Insert Table 8 Here

.....

Another way to evaluate model differences at the item level was to use a generalized MIRT model, or a reparameterization of both the compensatory and noncompensatory models into a single MIRT model, or

$$P_j(\theta_i) = c_j + (1-c_j) \frac{e^{f_{ij1} \cdot f_{ij2}}}{1 + e^{f_{ij1} \cdot f_{ij2}} + \mu(e^{f_{ij1}} + e^{f_{ij2}})}, \quad (6)$$

where  $\mu$  represented an indicator variable such that

$$\mu = \begin{cases} 0, & \text{for the linear or compensatory MIRT model,} \\ 1, & \text{for the multiplicative or noncompensatory MIRT model.} \end{cases}$$

Item response data,  $x_{ij}$ , were generated from samples of size  $N = 2000$  of  $\theta_i$  drawn from the bivariate normal distributions mentioned previously. The response data were known to have been produced by either the compensatory or noncompensatory MIRT model and were simulated by comparing the known values of  $P_j(\theta_i)$  to a pseudorandomly drawn uniform deviate,  $\omega$ , such that

$$x_{ij} = \begin{cases} 1, & 0 \leq \omega < P_j(\theta_i) \\ 0, & P_j(\theta_i) \leq \omega < 1. \end{cases}$$

The least squares estimation procedure was used to estimate the generalized MIRT model parameters. Each estimation was replicated 10 times with randomly selected starting values. Either four or five unique item parameters were estimated from the generalized MIRT model, as given by equation (6). The same item parameters that were given in tables 1-5 were used to generate the response data for the estimation procedure. When the response data were generated by the compensatory model,  $a_1$ ,  $a_2$ , and  $d$  (i.e.,  $d = -a_1b_1 - a_2b_2$ ) as well as  $\mu$ , were estimated. When the response data were generated by the noncompensatory model,  $a_1$ ,  $a_2$ ,  $b_1$ ,  $b_2$ , and  $\mu$  were estimated.

Table 9 shows the average bias in the item parameter estimates and the standard deviations of the estimates (in parentheses). For compensatory data, the model parameter,  $\mu$  was estimated fairly accurately for the uncorrelated situation, but the amount of bias and the standard deviation of the estimates increased as  $\rho$  increased. A similar situation occurred with noncompensatory data. However, although the amount of estimation error increased as the correlation between the abilities increased, the model still remained *identifiable*, in the sense that for compensatory data, the  $\mu$  estimates were statistically "close" to zero.



Likewise, for noncompensatory data, the  $\mu$  estimates were statistically "close" to one.

---

Insert Table 9 Here

---

The IOI analysis and the generalized MIRT model estimation gave similar results. That is, there were model differences at the item level, but these differences tended to decrease as the correlation in abilities increased. The generalized MIRT analysis also suggested that these differences might still be estimable, however, even when abilities are strongly correlated.

### Summary and Conclusions

These analyses and results seem to indicate that even though it is difficult to observe model differences at the overall test score level, there still may be measurable differences between the responses at the item level. Because the matching criterion between the two models resulted in similar expected  $p$ -values, we anticipated small differences at the total test score response level, or at the true score level. The differences that were detected at this level were consistent with the differences implied in the two models. Fewer high, number-correct scores or estimated true scores were observed from the noncompensatory model, but these and other total test differences decreased as  $\rho$  increased. As for the item response level analysis, both the IOI and the generalized MIRT model estimation showed that it is possible to quantify these differences and to distinguish between the data generated by carefully matched item response models of these two types. However, these differences, although real, are very small and probably not significant from any practical standpoint.

Although it is difficult to generalize beyond the two-dimensional situation used in the present study, it would appear to be difficult to distinguish between the two models without the benefit of any prior knowledge of item parameters or abilities. Even *with* such prior knowledge, response data generated by the models are nearly indistinguishable, especially with correlated abilities, which is likely the case in many real testing situations.

## References

- Ansley, T. N., & Forsyth, R. A. (1985). An examination of the characteristics of unidimensional IRT parameter estimates derived from two-dimensional data. *Applied Psychological Measurement*, 9, 37-48.
- Davey, T., Levine, M., & Williams, B. (1989, May). *Quantifying item and model fit*. Paper presented at the Office of Naval Research Contractor's Meeting on Model-based Psychological Measurement, Norman, OK.
- Doddy-Bogan, E., & Yen, W. M. (1983, April). *Detecting multidimensionality and examining its effects on vertical equating with the three-parameter logistic model*. Paper presented at the Annual meeting of the American Educational Research Association, Montreal.
- Embretson, S. (1984). A general latent trait model for response processes. *Psychometrika*, 49, 175-186.
- Hattie, J. (1981). *Decision criteria for determining unidimensionality*. Unpublished doctoral dissertation, University of Toronto, Toronto.
- Levine, M. V., Drasgow, F., Williams, B., McCusker, C., & Thomasson, G. (1990). *Distinguishing between item response models*. Unpublished manuscript.
- Lord, F. M., & Wingersky, M. S. (1984). Comparison of IRT true-score and equipercentile observed-score "equatings." *Applied Psychological Measurement*, 8, 453-461.
- Reckase, M. D. (1985). The difficulty of test items that measure more than one ability. *Applied Psychological Measurement*, 9, 401-412.
- Reckase, M. D. (1986, April). *The discriminating power of items that measure more than one dimension*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- Samejima, F. (1974). Normal ogive model on the continuous response level in the multidimensional latent space. *Psychometrika*, 39, 111-121.

### References

- Sympson, J. B. (1978). A model for testing with multidimensional items. In D. J. Weiss (Ed.), *Proceedings of the 1977 Computerized Adaptive Testing Conference* (pp. 82-98). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program.

## Appendix

### Analytical Definition of the *Ideal Observer Index*

A hypothetical observer is presented with two abilities,  $t_1$  and  $t_2$ , each with their associated item responses,  $u_1$  and  $u_2$ . The observer is informed that one ability-response pair was generated by one of two competing item response models, while the other pair was generated under the second model. The task is to correctly match each ability-response pair with the proper generating model. To make this decision, the observer is given access to both competing item response functions,  $P_1$  and  $P_2$ , and the common ability distribution,  $f(t)$ .

An ideal observer bases this decision on an optimal rule,  $\delta$ , which is determined by the ratio of likelihood functions,  $L_i(t_j, u_j) = P_i(t_j)^{u_j} Q_i(t_j)^{1-u_j}$ , where  $Q_i(t_j) = 1 - P_i(t_j)$ ,  $i = 1, 2$ ;  $j = 1, 2$ . The decision rule,  $\delta$ , is then defined as

$$\delta = \begin{array}{l} \text{if } L_1(t_1, u_1) \cdot L_2(t_2, u_2) > L_1(t_2, u_2) \cdot L_2(t_1, u_1), \text{ then decide model } \{P_1; f\} \\ \text{produced sample } \{t_1, u_1\} \text{ while model } \{P_2; f\} \text{ produced } \{t_2, u_2\}. \\ \\ \text{if } L_1(t_2, u_2) \cdot L_2(t_1, u_1) > L_1(t_1, u_1) \cdot L_2(t_2, u_2), \text{ then decide model } \{P_2; f\} \\ \text{produced sample } \{t_1, u_1\} \text{ while model } \{P_1; f\} \text{ produced } \{t_2, u_2\}. \end{array}$$

The probability of this decision rule being correct, given the model, is

$$\begin{aligned} \text{Prob}[\delta \text{ correct} | \text{model}] = & \text{Prob}[L_1(t_1, u_1) \cdot L_2(t_2, u_2) > L_1(t_2, u_2) \cdot L_2(t_1, u_1) | \{P_1; f\} \& \{P_2; f\}] + \\ & \text{Prob}[L_1(t_2, u_2) \cdot L_2(t_1, u_1) > L_1(t_1, u_1) \cdot L_2(t_2, u_2) | \{P_2; f\} \& \{P_1; f\}]. \end{aligned}$$

The response pair,  $\mathbf{u}$ , where  $\mathbf{u} = (u_1, u_2)$ , can be defined in four possible patterns: (1,1), (1,0), (0,1), and (0,0). Therefore,

$$\begin{aligned} & \text{Prob}[L_1(t_1, u_1) \cdot L_2(t_2, u_2) > L_1(t_2, u_2) \cdot L_2(t_1, u_1) | \{P_1; f\} \& \{P_2; f\}] = \\ & \text{Prob}[P_1(t_1) \cdot P_2(t_2) > P_1(t_2) \cdot P_2(t_1) | \mathbf{u} = (1,1)] \cdot \text{Prob}[\mathbf{u} = (1,1) | \{P_1; f\} \& \{P_2; f\}] \\ & + \text{Prob}[P_1(t_1) \cdot Q_2(t_2) > Q_1(t_2) \cdot P_2(t_1) | \mathbf{u} = (1,0)] \cdot \text{Prob}[\mathbf{u} = (1,0) | \{P_1; f\} \& \{P_2; f\}] \\ & + \text{Prob}[Q_1(t_1) \cdot P_2(t_2) > P_1(t_2) \cdot Q_2(t_1) | \mathbf{u} = (0,1)] \cdot \text{Prob}[\mathbf{u} = (0,1) | \{P_1; f\} \& \{P_2; f\}] \\ & + \text{Prob}[Q_1(t_1) \cdot Q_2(t_2) > Q_1(t_2) \cdot Q_2(t_1) | \mathbf{u} = (0,0)] \cdot \text{Prob}[\mathbf{u} = (0,0) | \{P_1; f\} \& \{P_2; f\}]. \end{aligned}$$

$$\text{Define } \pi_{ij} = \iint P_1(t)^{u_i} Q_1(t)^{1-u_i} P_2(g)^{u_j} Q_2(g)^{1-u_j} f(t) f(g) dt dg.$$

$$\text{Then, } \text{Prob}[L_1(t_1, u_1) \cdot L_2(t_2, u_2) > L_1(t_2, u_2) \cdot L_2(t_1, u_1) | \{P_1; f\} \& \{P_2; f\}] =$$

$$\begin{aligned} & \pi_{11} \text{Prob}[P_1(t_1) \cdot P_2(t_2) > P_1(t_2) \cdot P_2(t_1) | \mathbf{u} = (1,1)] + \\ & \pi_{10} \text{Prob}[P_1(t_1) \cdot Q_2(t_2) > Q_1(t_2) \cdot P_2(t_1) | \mathbf{u} = (1,0)] + \\ & \pi_{01} \text{Prob}[Q_1(t_1) \cdot P_2(t_2) > P_1(t_2) \cdot Q_2(t_1) | \mathbf{u} = (0,1)] + \\ & \pi_{00} \text{Prob}[Q_1(t_1) \cdot Q_2(t_2) > Q_1(t_2) \cdot Q_2(t_1) | \mathbf{u} = (0,0)]. \end{aligned}$$

$$\text{Similarly, } \text{Prob}[L_1(t_2, u_2) \cdot L_2(t_1, u_1) > L_1(t_1, u_1) \cdot L_2(t_2, u_2) | \{P_2; f\} \& \{P_1; f\}] =$$

$$\begin{aligned} & \text{Prob}[P_1(t_2) \cdot P_2(t_1) > P_1(t_1) \cdot P_2(t_2) | \mathbf{u} = (1,1)] \cdot \text{Prob}[\mathbf{u} = (1,1) | \{P_2; f\} \& \{P_1; f\}] \\ & + \text{Prob}[P_1(t_2) \cdot Q_2(t_1) > Q_1(t_1) \cdot P_2(t_2) | \mathbf{u} = (1,0)] \cdot \text{Prob}[\mathbf{u} = (1,0) | \{P_2; f\} \& \{P_1; f\}] \\ & + \text{Prob}[Q_1(t_2) \cdot P_2(t_1) > P_1(t_1) \cdot Q_2(t_2) | \mathbf{u} = (0,1)] \cdot \text{Prob}[\mathbf{u} = (0,1) | \{P_2; f\} \& \{P_1; f\}] \\ & + \text{Prob}[Q_1(t_2) \cdot Q_2(t_1) > Q_1(t_1) \cdot Q_2(t_2) | \mathbf{u} = (0,0)] \cdot \text{Prob}[\mathbf{u} = (0,0) | \{P_2; f\} \& \{P_1; f\}]. \end{aligned}$$

Then,  $\text{Prob}[L_1(t_2, u_2) \cdot L_2(t_1, u_1) > L_1(t_1, u_1) \cdot L_2(t_2, u_2) | \{P_2; f\} \& \{P_1; f\}] =$

$$\begin{aligned} & \pi_{11} \text{Prob}[P_1(t_2) \cdot P_2(t_1) > P_1(t_1) \cdot P_2(t_2) | u = (1,1)] + \\ & \pi_{10} \text{Prob}[P_1(t_2) \cdot Q_2(t_1) > Q_1(t_1) \cdot P_2(t_2) | u = (1,0)] + \\ & \pi_{01} \text{Prob}[Q_1(t_2) \cdot P_2(t_1) > P_1(t_1) \cdot Q_2(t_2) | u = (0,1)] + \\ & \pi_{00} \text{Prob}[Q_1(t_2) \cdot Q_2(t_1) > Q_1(t_1) \cdot Q_2(t_2) | u = (0,0)]. \end{aligned}$$

Let  $\Omega_{u_1 u_2}$  be defined as that region of the ability space where

$$P_1(t_1)^{u_1} \cdot Q_1(t_1)^{u_1-1} \cdot P_2(t_2)^{u_2} \cdot Q_2(t_2)^{u_2-1} > P_1(t_2)^{u_2} \cdot Q_1(t_2)^{u_2-1} \cdot P_2(t_1)^{u_1} \cdot Q_2(t_1)^{u_1-1}$$

holds, and likewise let  $\bar{\Omega}_{u_1 u_2}$  be defined as that region of the ability space where

$$P_1(t_2)^{u_2} \cdot Q_1(t_2)^{u_2-1} \cdot P_2(t_1)^{u_1} \cdot Q_2(t_1)^{u_1-1} > P_1(t_1)^{u_1} \cdot Q_1(t_1)^{u_1-1} \cdot P_2(t_2)^{u_2} \cdot Q_2(t_2)^{u_2-1}$$

is true. Then

$$\text{Prob}[P_1(t_1) \cdot P_2(t_2) > P_1(t_2) \cdot P_2(t_1) | u = (1,1)] = \int \int_{\Omega_{11}} f(t) f(g) dt dg,$$

$$\text{Prob}[P_1(t_1) \cdot Q_2(t_2) > Q_1(t_2) \cdot P_2(t_1) | u = (1,0)] = \int \int_{\Omega_{10}} f(t) f(g) dt dg,$$

$$\text{Prob}[Q_1(t_1) \cdot P_2(t_2) > P_1(t_2) \cdot Q_2(t_1) | u = (0,1)] = \int \int_{\Omega_{01}} f(t) f(g) dt dg,$$

and

$$\text{Prob}[Q_1(t_1) \cdot Q_2(t_2) > Q_1(t_2) \cdot Q_2(t_1) | u = (0,0)] = \int \int_{\Omega_{00}} f(t) f(g) dt dg.$$

Then

$$\text{Prob}[P_1(t_2) \cdot P_2(t_1) > P_1(t_1) \cdot P_2(t_2) | \mathbf{u} = (1,1)] = \int \int_{\mathfrak{a}_{11}} f(t) f(g) dt dg,$$

$$\text{Prob}[P_1(t_2) \cdot Q_2(t_1) > Q_1(t_1) \cdot P_2(t_2) | \mathbf{u} = (0,1)] = \int \int_{\mathfrak{a}_{01}} f(t) f(g) dt dg,$$

$$\text{Prob}[Q_1(t_2) \cdot P_2(t_1) > Q_2(t_2) \cdot P_1(t_1) | \mathbf{u} = (1,0)] = \int \int_{\mathfrak{a}_{10}} f(t) f(g) dt dg,$$

and

$$\text{Prob}[Q_1(t_2) \cdot Q_2(t_1) > Q_1(t_1) \cdot Q_2(t_2) | \mathbf{u} = (0,0)] = \int \int_{\mathfrak{a}_{00}} f(t) f(g) dt dg.$$

Thus,  $\text{Prob}[\delta \text{ correct} | \text{model}] =$

$$\begin{aligned} & \pi_{11} \int \int_{\mathfrak{a}_{11}} f(t) f(g) dt dg + \pi_{10} \int \int_{\mathfrak{a}_{10}} f(t) f(g) dt dg + \\ & \pi_{01} \int \int_{\mathfrak{a}_{01}} f(t) f(g) dt dg + \pi_{00} \int \int_{\mathfrak{a}_{00}} f(t) f(g) dt dg + \\ & \pi_{11} \int \int_{\mathfrak{a}_{11}} f(t) f(g) dt dg + \pi_{10} \int \int_{\mathfrak{a}_{01}} f(t) f(g) dt dg + \\ & \pi_{01} \int \int_{\mathfrak{a}_{10}} f(t) f(g) dt dg + \pi_{00} \int \int_{\mathfrak{a}_{00}} f(t) f(g) dt dg \end{aligned}$$

or

$$\begin{aligned} & \pi_{11} + \pi_{10} \left( \int \int_{\mathfrak{a}_{10}} f(t) f(g) dt dg + \int \int_{\mathfrak{a}_{01}} f(t) f(g) dt dg \right) + \\ & \pi_{01} \left( \int \int_{\mathfrak{a}_{01}} f(t) f(g) dt dg + \int \int_{\mathfrak{a}_{10}} f(t) f(g) dt dg \right) + \pi_{00} \end{aligned}$$



Finally,  $\text{Prob}[\delta \text{ correct}] = \text{Prob}[\delta \text{ correct} | \text{model}] \cdot \text{Prob}[\text{selecting a model}]$ . Because each model is equally likely, the probability of selecting a model is equal to .5. Thus,  $\text{Prob}[\delta \text{ correct}] = .5(\text{Prob}[\delta \text{ correct} | \text{model}])$ .

Table 1

*Original Item Parameters for the Compensatory Model*

Item #	$a_1$	$a_2$	d	E(p-value) rho			
				.00	.25	.50	.75
01	0.90	1.31	-0.67	.39	.39	.40	.40
02	2.10	0.50	-1.13	.34	.35	.36	.36
03	0.89	1.10	0.52	.59	.59	.58	.58
04	0.99	1.00	-0.44	.42	.42	.43	.43
05	0.58	1.65	0.78	.63	.62	.62	.61
06	0.91	1.27	0.42	.57	.57	.57	.56
07	1.03	0.95	1.08	.69	.68	.67	.67
08	0.32	2.27	0.38	.55	.55	.55	.55
09	0.61	0.72	1.63	.80	.79	.79	.78
10	0.67	1.12	0.60	.61	.61	.60	.60
11	0.91	0.91	-0.21	.46	.46	.46	.47
12	0.64	1.72	-0.05	.49	.49	.49	.49
13	1.65	0.38	0.40	.57	.56	.56	.56
14	0.18	1.61	1.84	.78	.78	.78	.77
15	0.82	1.02	0.09	.52	.52	.52	.51
16	1.45	0.81	-0.24	.46	.46	.46	.46
17	1.64	0.62	0.85	.64	.63	.63	.62
18	0.77	0.76	-0.91	.32	.33	.34	.34
19	1.46	0.62	0.10	.52	.52	.52	.52
20	0.39	1.37	0.32	.56	.56	.55	.55

Table 2

*Item Parameters for the Noncompensatory Model with  $Rho = .00$* 

Item #	$a_1$	$a_2$	$b_1$	$b_2$	E(p-value)
01	1.26	1.60	-0.92	-0.15	.38
02	2.30	1.04	0.38	-2.28	.34
03	1.22	1.39	-1.42	-0.99	.59
04	1.32	1.35	-0.62	-0.58	.42
05	1.02	1.82	-2.71	-0.62	.62
06	1.25	1.53	-1.45	-0.79	.56
07	1.30	1.26	-1.48	-1.63	.68
08	0.92	2.38	-3.95	-0.22	.55
09	0.93	1.00	-2.75	-2.35	.80
10	1.05	1.37	-1.96	-0.90	.61
11	1.24	1.25	-0.78	-0.75	.46
12	1.07	1.92	-2.17	-0.19	.49
13	1.81	0.88	-0.36	-3.25	.56
14	0.85	1.67	-5.26	-1.17	.78
15	1.17	1.32	-1.21	-0.75	.51
16	1.71	1.23	-0.27	-1.35	.45
17	1.83	1.06	-0.6	-2.55	.63
18	1.09	1.09	-0.31	-0.32	.32
19	1.69	1.07	-0.35	-1.98	.51
20	0.88	1.54	-2.98	-0.41	.55

Table 3

*Item Parameters for the Noncompensatory Model with  $Rho = .25$*

Item #	$a_1$	$a_2$	$b_1$	$b_2$	E(p-value)
01	1.38	1.74	-0.79	-0.14	.39
02	2.40	1.14	0.35	-1.88	.34
03	1.36	1.50	-1.27	-0.91	.58
04	1.44	1.45	-0.56	-0.51	.42
05	1.17	1.94	-2.30	-0.60	.61
06	1.40	1.66	-1.28	-0.73	.56
07	1.45	1.40	-1.34	-1.47	.72
08	1.05	2.47	-3.30	-0.22	.55
09	1.02	1.09	-2.49	-2.17	.79
10	1.17	1.47	-1.72	-0.85	.60
11	1.34	1.34	-0.71	-0.68	.46
12	1.21	2.06	-1.82	-0.20	.49
13	1.90	0.98	-0.36	-2.80	.56
14	0.93	1.72	-4.65	-1.15	.78
15	1.29	1.42	-1.08	-0.69	.51
16	1.84	1.33	-0.27	-1.16	.45
17	1.97	1.20	-0.66	-2.19	.62
18	1.15	1.16	-0.28	-0.27	.33
19	1.80	1.18	-0.35	-1.71	.51
20	0.98	1.61	-2.57	-0.40	.55

Table 4

*Item Parameters for the Noncompensatory Model with  $Rho = .50$*

Item #	$a_1$	$a_2$	$b_1$	$b_2$	E(p-value)
01	1.52	1.82	-0.66	-0.12	.39
02	2.48	1.27	0.32	-1.51	.35
03	1.49	1.63	-1.14	-0.85	.58
04	1.54	1.54	-0.50	-0.45	.42
05	1.32	2.04	-1.97	-0.59	.61
06	1.55	1.79	-1.13	-0.68	.56
07	1.58	1.55	-1.23	-1.33	.67
08	1.20	2.51	-2.78	-0.22	.55
09	1.10	1.17	-2.30	-2.03	.78
10	1.28	1.56	-1.53	-0.80	.60
11	1.44	1.43	-0.64	-0.61	.46
12	1.36	2.13	-1.54	-0.19	.49
13	1.96	1.09	-0.36	-2.39	.56
14	1.03	1.77	-4.07	-1.13	.77
15	1.39	1.51	-0.97	-0.64	.51
16	1.95	1.47	-0.26	-0.99	.46
17	2.08	1.35	-0.63	-1.89	.62
18	1.21	1.20	-0.23	-0.23	.33
19	1.89	1.30	-0.34	-1.46	.51
20	1.08	1.66	-2.23	-0.40	.55

Table 5

*Item Parameters for the Noncompensatory Model with  $Rho = .75$* 

Item #	$a_1$	$a_2$	$b_1$	$b_2$	E(p-value)
01	1.65	1.92	-0.51	-0.10	.40
02	2.53	1.43	0.31	-1.14	.35
03	1.60	1.73	-1.01	-0.77	.58
04	1.63	1.64	-0.42	-0.39	.43
05	1.48	2.14	-1.67	-0.57	.61
06	1.69	1.92	-0.98	-0.62	.56
07	1.69	1.66	-1.13	-1.21	.66
08	1.36	2.57	-2.25	-0.22	.55
09	1.15	1.21	-2.17	-1.93	.78
10	1.38	1.63	-1.36	-0.76	.60
11	1.50	1.51	-0.56	-0.54	.46
12	1.53	2.23	-1.22	-0.19	.49
13	1.98	1.26	-0.36	-1.99	.56
14	1.15	1.78	-3.60	-1.11	.77
15	1.47	1.59	-0.85	-0.59	.51
16	2.03	1.63	-0.23	-0.81	.46
17	2.15	1.53	-0.61	-1.60	.62
18	1.24	1.24	-0.18	-0.18	.34
19	1.94	1.44	-0.33	-1.23	.51
20	1.17	1.70	-1.92	-0.40	.55

Table 6

*Compensatory Minus Noncompensatory Density Differences in Number-correct Score*

Number-correct score (y)	rho			
	.00	.25	.50	.75
20	.013	.014	.014	.011
19	.015	.012	.009	.004
18	.012	.007	.003	.000
17	.007	.003	.000	-.002
16	.002	-.001	-.002	-.003
15	-.003	-.003	-.004	-.003
14	-.006	-.005	-.004	-.003
13	-.009	-.007	-.005	-.003
12	-.011	-.007	-.005	-.002
11	-.012	-.008	-.005	-.002
10	-.012	-.008	-.004	-.001
9	-.011	-.007	-.004	-.001
8	-.009	-.006	-.003	.000
7	-.006	-.004	-.002	.001
6	-.003	-.002	-.001	.001
5	.001	.000	.001	.002
4	.005	.002	.002	.002
3	.008	.005	.003	.002
2	.009	.006	.004	.001
1	.008	.006	.003	-.001
0	-.005	.003	.001	-.004

Table 7

*Central Moments of Number-correct Scores*

MIRT Models	rho	Mean	Second Central Moment	Third Central Moment	Fourth Central Moment
Compensatory	.00	10.90	25.79	-16.56	1362.83
	.25	10.88	29.40	-20.44	1680.36
	.50	10.86	32.64	-24.01	1980.03
	.75	10.84	35.57	-27.27	2262.98
Noncompensatory	.00	10.79	20.67	-9.42	946.49
	.25	10.78	25.43	-15.86	1336.75
	.50	10.78	30.12	-24.30	1760.64
	.75	10.78	34.70	-32.74	2200.57



Table 8

*Ideal Observer Index*

Item #	rho			
	.00	.25	.50	.75
01	.5479	.5397	.5295	.5179
02	.5311	.5265	.5205	.5128
03	.5513	.5418	.5307	.5183
04	.5461	.5377	.5279	.5171
05	.5421	.5353	.5265	.5157
06	.5550	.5451	.5332	.5194
07	.5511	.5419	.5304	.5175
08	.5243	.5212	.5165	.5102
09	.5276	.5227	.5162	.5092
10	.5430	.5351	.5254	.5149
11	.5435	.5355	.5260	.5156
12	.5448	.5375	.5281	.5166
13	.5291	.5246	.5185	.5112
14	.5124	.5109	.5082	.5048
15	.5456	.5370	.5271	.5161
16	.5497	.5411	.5307	.5182
17	.5442	.5371	.5276	.5162
18	.5281	.5232	.5175	.5114
19	.5425	.5352	.5260	.5156
20	.5292	.5241	.5179	.5108

Table 9

*Average Bias (parameter estimate - true parameter) and Standard Deviation  
of Bias in Estimates of the Generalized MIRT Model Parameters*

Response Data Model	rho	a <sub>1</sub>	a <sub>2</sub>	d	b <sub>1</sub>	b <sub>2</sub>	μ
Compensatory	.00	.044 (.042)	.024 (.073)	.069 (.158)			.013 (.009)
	.25	.044 (.047)	.040 (.042)	.125 (.275)			.026 (.052)
	.50	.078 (.055)	.069 (.081)	.255 (.238)			.064 (.060)
	.75	.098 (.128)	.113 (.080)	.787 (1.930)			.107 (.094)
Noncompensatory	.00	-.008 (.099)	.009 (.115)		.130 (.448)	.230 (.354)	-.199 (.163)
	.25	-.006 (.090)	-.004 (.083)		.250 (.622)	.254 (.464)	-.197 (.144)
	.50	.039 (.145)	-.076 (.104)		.191 (.888)	.183 (.265)	-.200 (.125)
	.75	-.155 (.220)	-.059 (.105)		.071 (.439)	.250 (.421)	-.288 (.175)

Note: standard deviations are in parentheses

**Figure Captions**

*Figure 1.* Difference Between Compensatory and Noncompensatory True Scores:  
Rho = .00

*Figure 2.* Difference Between Compensatory and Noncompensatory True Scores:  
Rho = .25

*Figure 3.* Difference Between Compensatory and Noncompensatory True Scores:  
Rho = .50

*Figure 4.* Difference Between Compensatory and Noncompensatory True Scores:  
Rho = .75

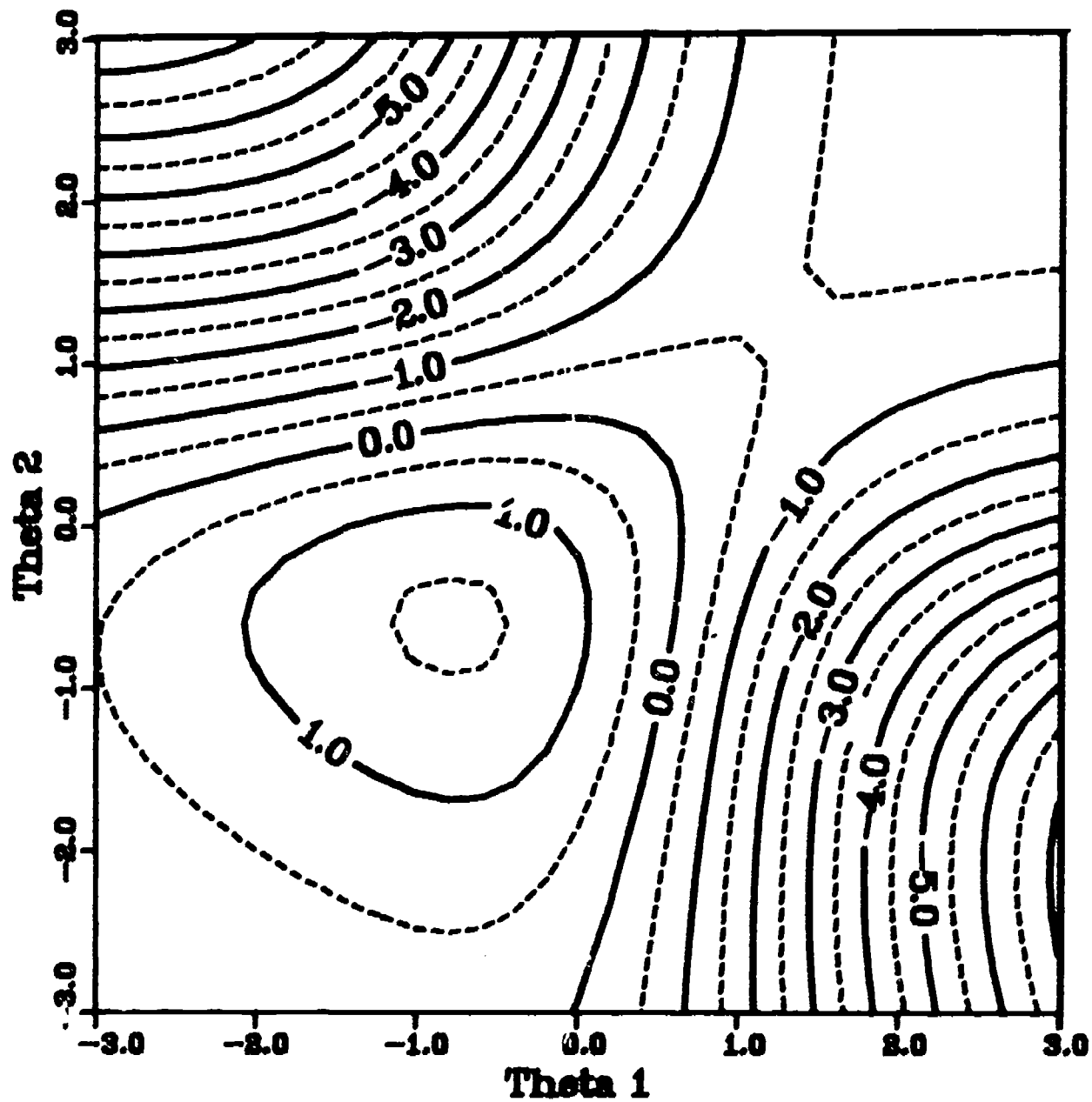
*Figure 5.* Absolute Difference Between Compensatory and Noncompensatory Test  
Information Vectors: Rho = .00

*Figure 6.* Absolute Difference Between Compensatory and Noncompensatory Test  
Information Vectors: Rho = .25

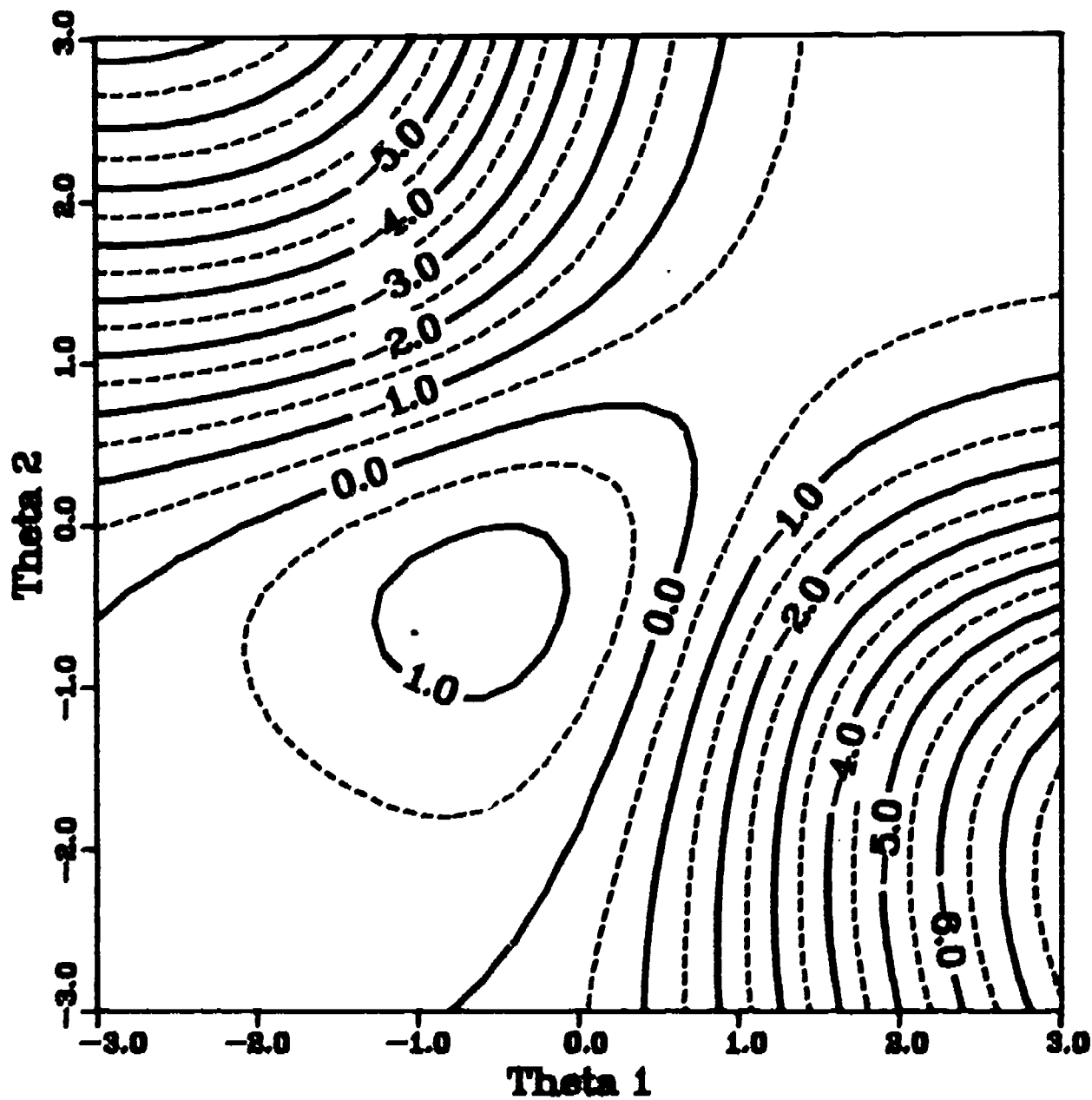
*Figure 7.* Absolute Difference Between Compensatory and Noncompensatory Test  
Information Vectors: Rho = .50

*Figure 8.* Absolute Difference Between Compensatory and Noncompensatory Test  
Information Vectors: Rho = .75

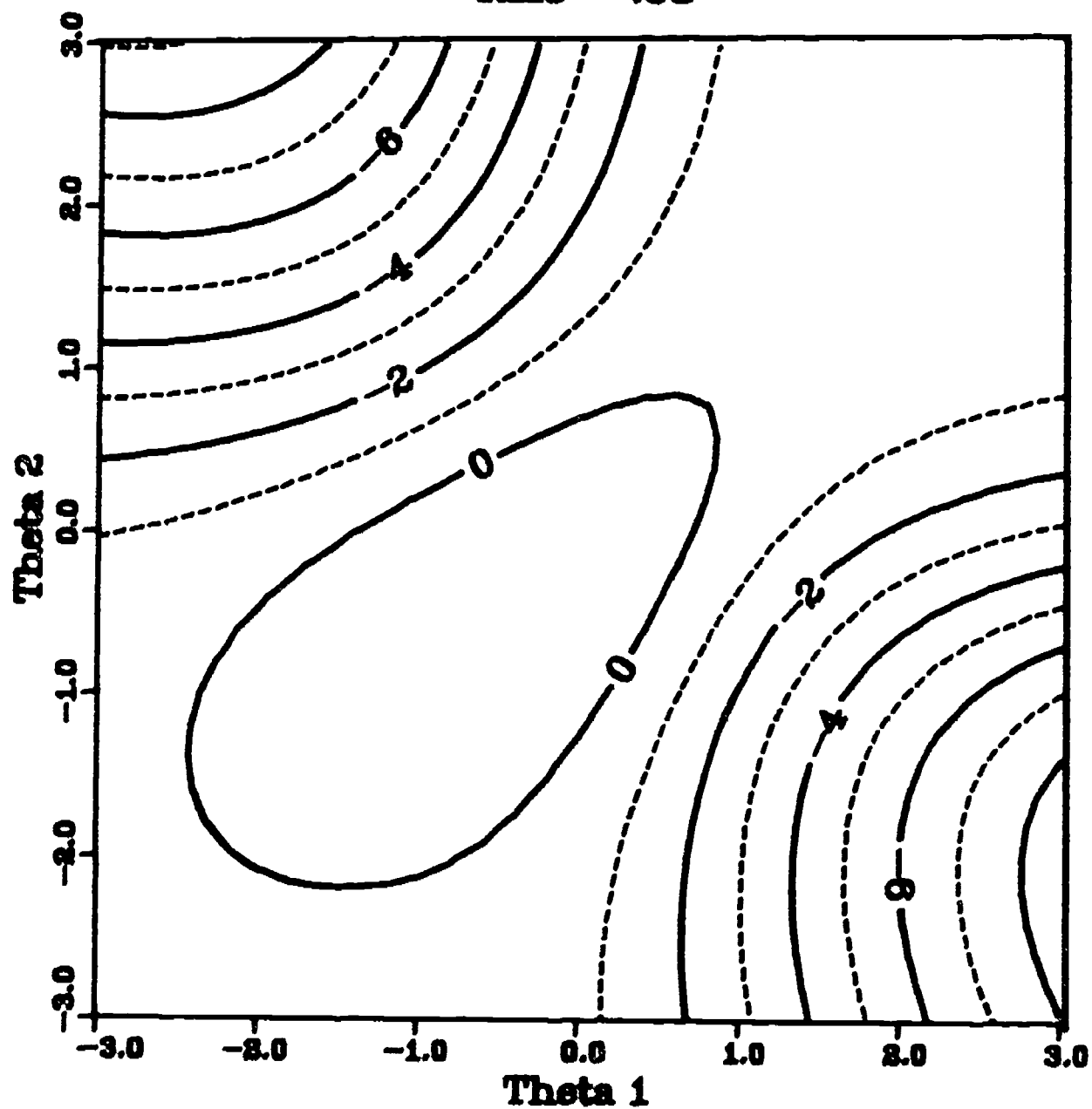
**Difference Between Compensatory  
and Noncompensatory True Scores:  
 $\text{Rho} = .00$**



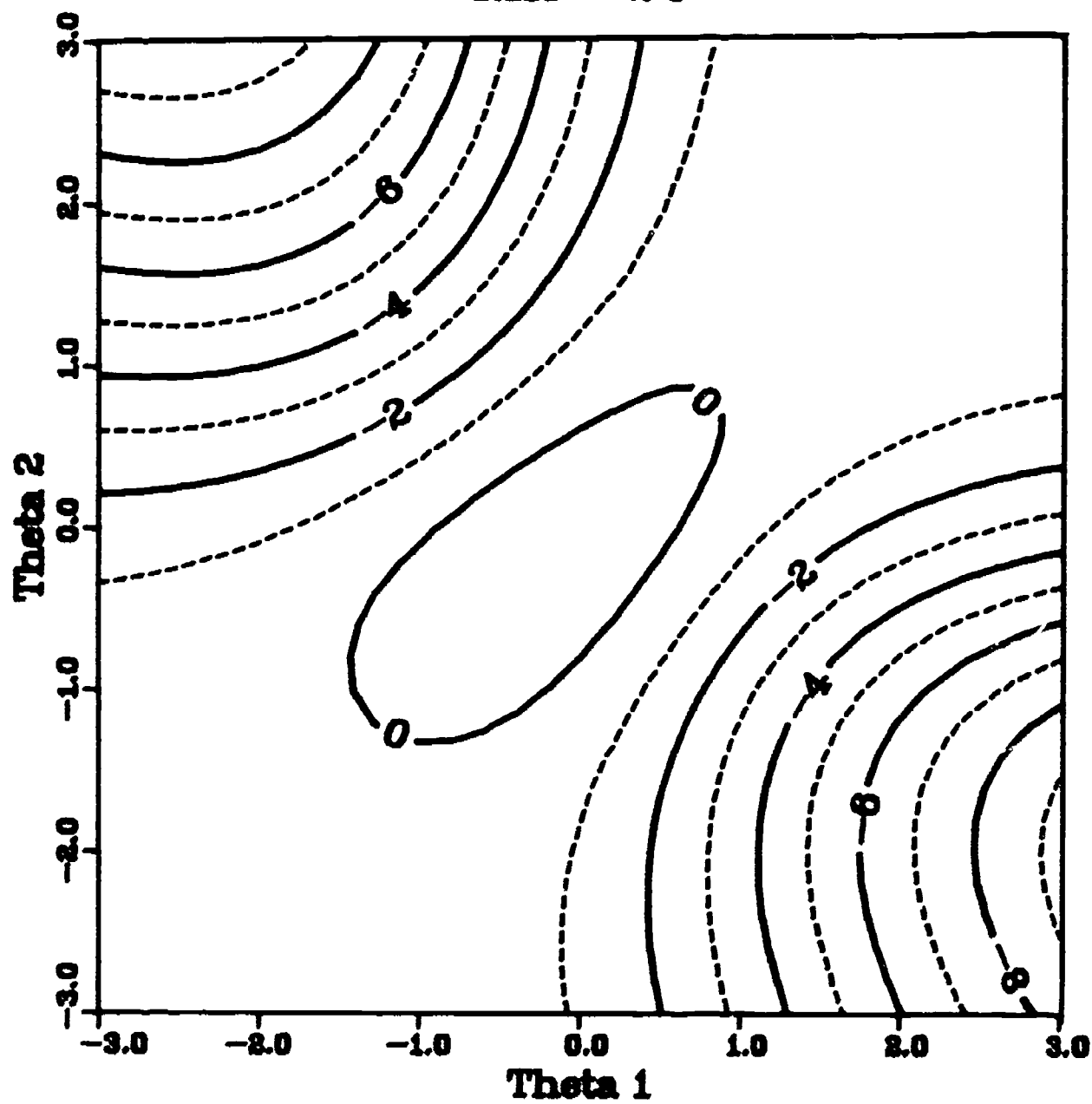
**Difference Between Compensatory  
and Noncompensatory True Scores:  
 $\text{Rho} = .25$**



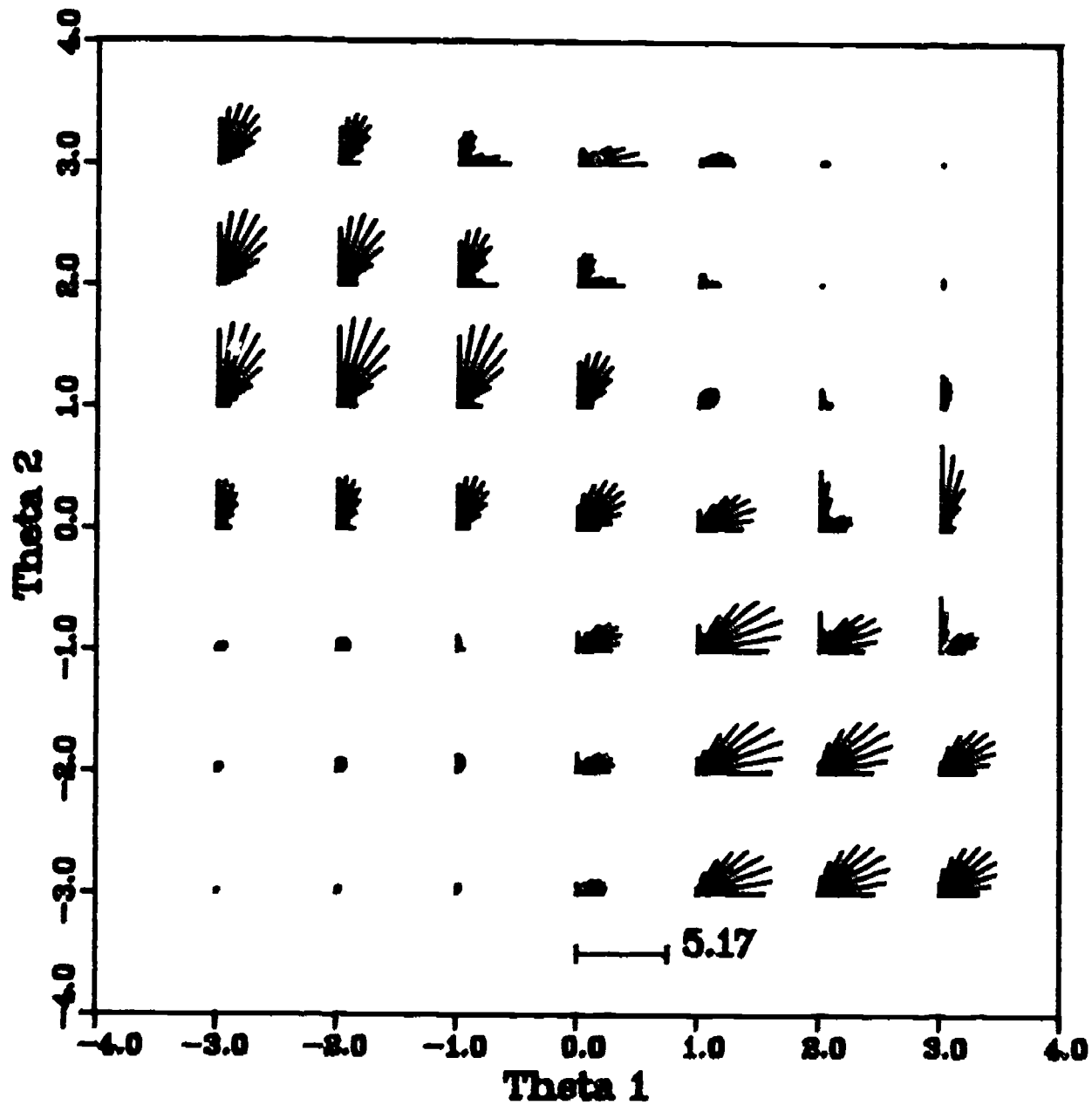
**Difference Between Compensatory  
and Noncompensatory True Scores:  
 $\text{Rho} = .50$**



**Difference Between Compensatory  
and Noncompensatory True Scores:  
 $\text{Rho} = .75$**

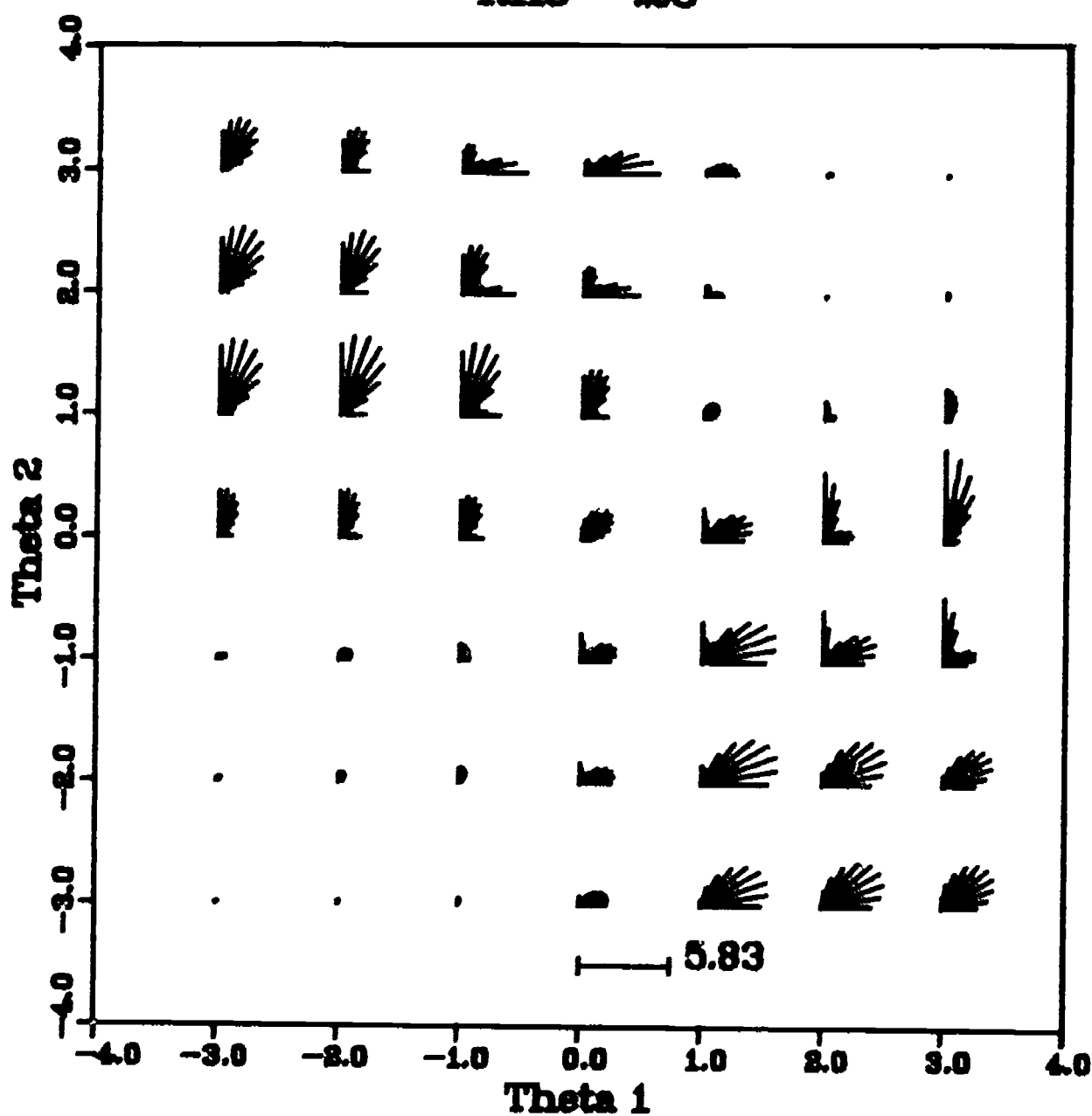


**Absolute Difference Between Compensatory  
and Noncompensatory Test Information Vectors:  
 $\text{Rho} = .00$**

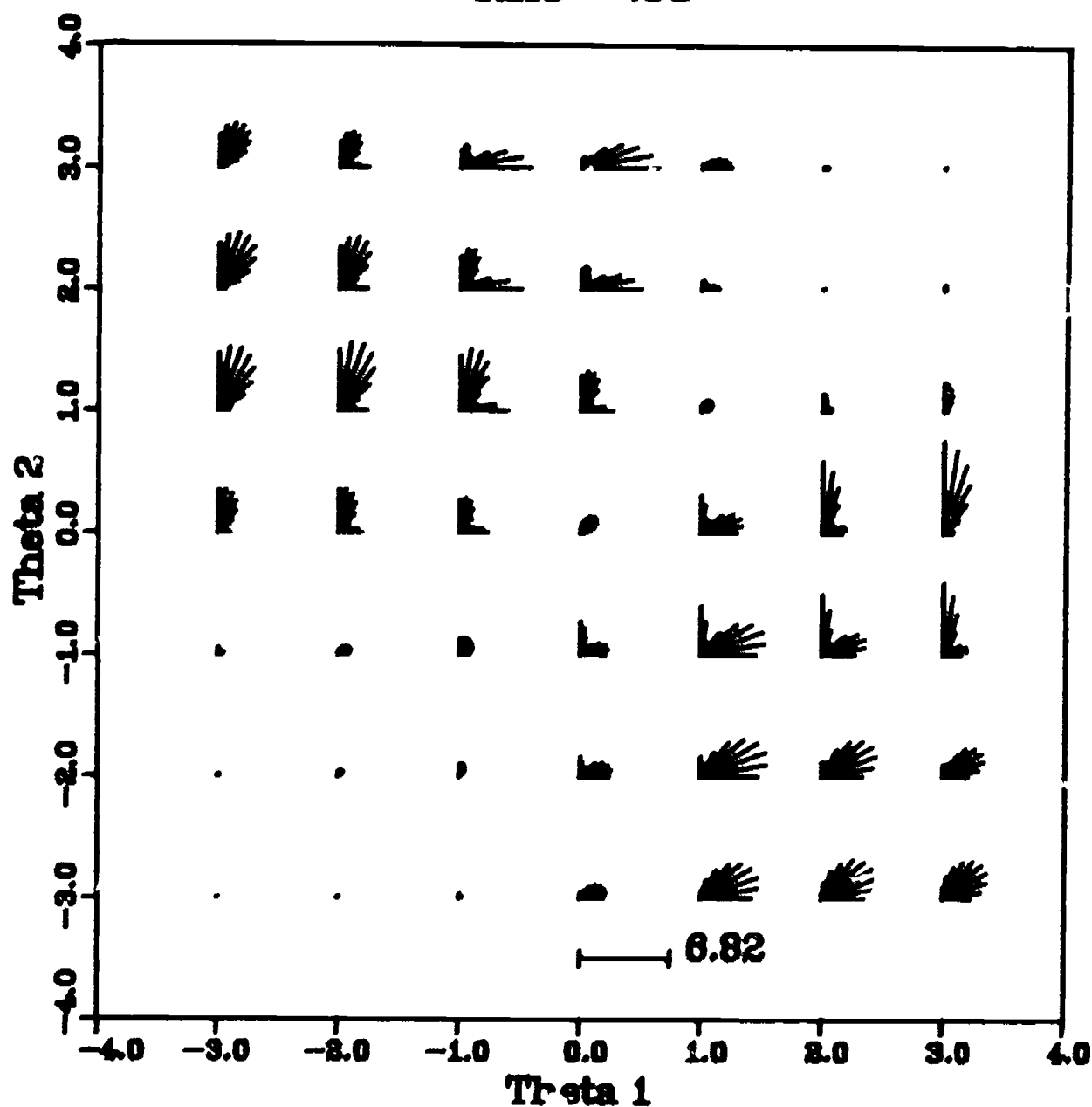




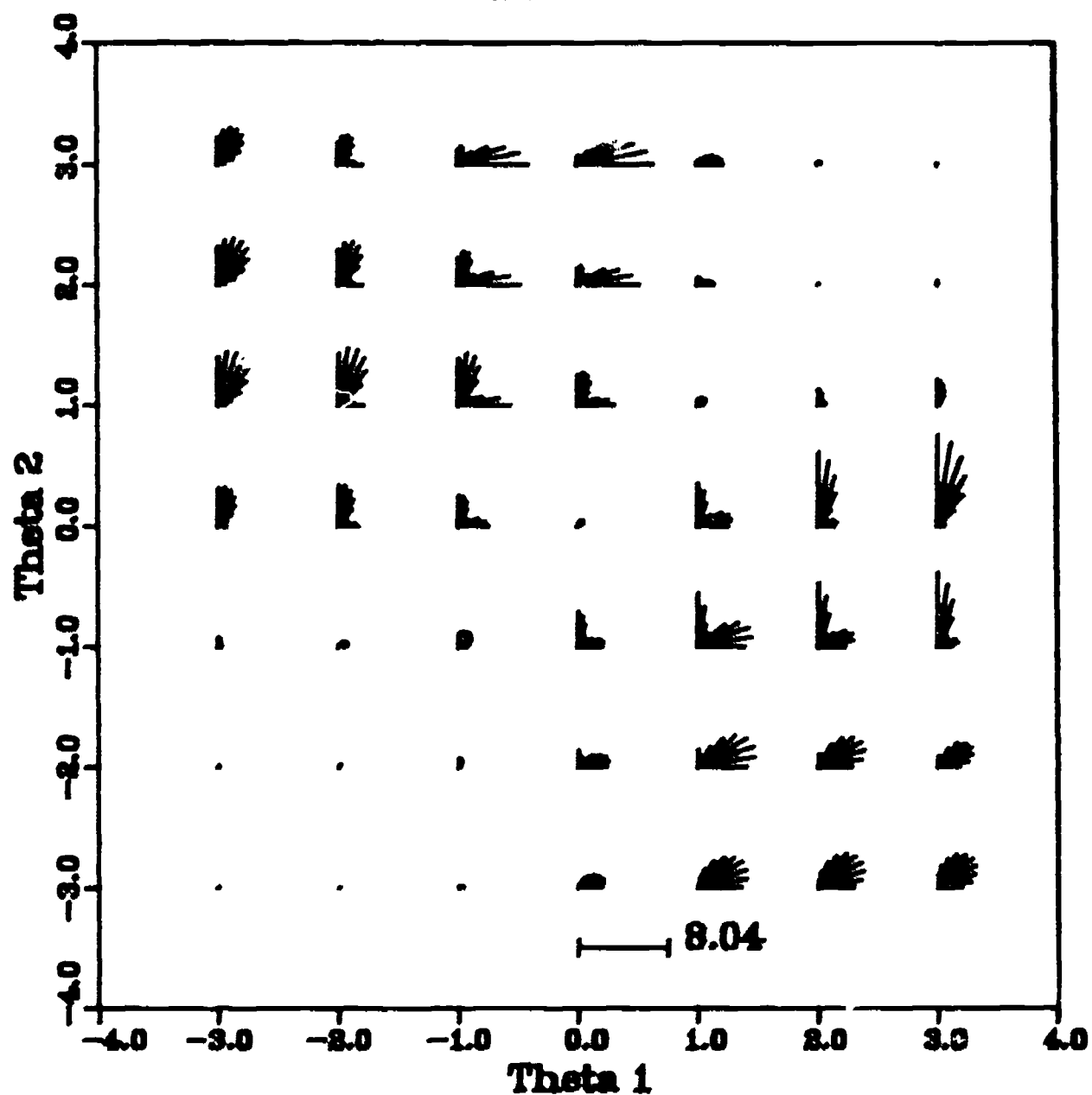
# **Absolute Difference Between Compensatory and Noncompensatory Test Information Vectors: $\rho = .25$**



**Absolute Difference Between Compensatory  
and Noncompensatory Test Information Vectors:  
 $\text{Rho} = .50$**



**Absolute Difference Between Compensatory  
and Noncompensatory Test Information Vectors:  
 $\text{Rho} = .75$**



Distribution List

Dr. Terry Ackerman  
Educational Psychology  
210 Education Bldg.  
University of Illinois  
Champaign, IL 61801

Dr. James Algina  
1403 Norman Hall  
University of Florida  
Gainesville, FL 32605

Dr. Erling B. Andersen  
Department of Statistics  
Studiestraße 6  
1455 Copenhagen  
DENMARK

Dr. Ronald Armstrong  
Rutgers University  
Graduate School of Management  
Newark, NJ 07102

Dr. Eva L. Baker  
UCLA Center for the Study  
of Evaluation  
145 Moore Hall  
University of California  
Los Angeles, CA 90024

Dr. Laura L. Barnes  
College of Education  
University of Toledo  
2801 W. Bancroft Street  
Toledo, OH 43606

Dr. William M. Bart  
University of Minnesota  
Dept. of Educ. Psychology  
330 Burton Hall  
178 Pillsbury Dr., S.E.  
Minneapolis, MN 55455

Dr. Isaac Bejar  
Law School Admissions  
Services  
P.O. Box 40  
Newtown, PA 18940-0040

Dr. Menucha Birenbaum  
School of Education  
Tel Aviv University  
Ramat Aviv 60078  
ISRAEL

Dr. Arthur S. Blumens  
Code N712  
Naval Training Systems Center  
Orlando, FL 32813-7100

Dr. Bruce Blumom  
Defense Manpower Data Center  
99 Pacific St.  
Suite 155A  
Monterey, CA 93943-3231

Cdt. Arnold Bohrer  
Sector Psychologisch Onderzoek  
Rekrutering En Selectiecentrum  
Kwartier Koningen Astrid  
Bruijnstraat  
1120 Brussels, BELGIUM

Dr. Robert Breaux  
Code 281  
Naval Training Systems Center  
Orlando, FL 32826-3224

Dr. Robert Brennan  
American College Testing  
Programs  
P. O. Box 168  
Iowa City, IA 52243

Dr. Gregory Candell  
CTB/McGraw-Hill  
2500 Garden Road  
Monterey, CA 93940

Dr. John B. Carroll  
409 Elliott Rd., North  
Chapel Hill, NC 27514

Dr. John M. Carroll  
IBM Watson Research Center  
User Interface Institute  
P.O. Box 704  
Yorktown Heights, NY 10598

Dr. Robert M. Carroll  
Chief of Naval Operations  
OP-0182  
Washington, DC 20350

Dr. Raymond E. Christal  
UES LAMP Science Advisor  
AFHRL/MOEL  
Brooks AFB, TX 78235

Mr. Hui Hui Chung  
University of Illinois  
Department of Statistics  
101 Illini Hall  
725 South Wright St.  
Champaign, IL 61820

Dr. Norman Cliff  
Department of Psychology  
Univ. of So. California  
Los Angeles, CA 90089-1061

Director, Manpower Program  
Center for Naval Analyses  
4401 Ford Avenue  
P.O. Box 16268  
Alexandria, VA 22302-0268

Director,  
Manpower Support and  
Readiness Program  
Center for Naval Analyses  
2000 North Beauregard Street  
Alexandria, VA 22311

Dr. Stanley Collier  
Office of Naval Technology  
Code 222  
800 N. Quincy Street  
Arlington, VA 22217-5000

Dr. Hans F. Crombag  
Faculty of Law  
University of Limburg  
P.O. Box 616  
Maastricht  
The NETHERLANDS 6200 MD

Ms. Carolyn R. Crone  
Johns Hopkins University  
Department of Psychology  
Charles & 34th Street  
Baltimore, MD 21218

Dr. Timothy Davey  
American College Testing Program  
P.O. Box 168  
Iowa City, IA 52243

Dr. C. M. Dayton  
Department of Measurement  
Statistics & Evaluation  
College of Education  
University of Maryland  
College Park, MD 20742

Dr. Ralph J. DeAvala  
Measurement, Statistics  
and Evaluation  
Benjamin Bldg., Rm. 4112  
University of Maryland  
College Park, MD 20742

Dr. Lou DiBello  
CERL  
University of Illinois  
103 South Mathews Avenue  
Urbana, IL 61801

Dr. Dattaprasad Divya  
Center for Naval Analyses  
4401 Ford Avenue  
P.O. Box 16268  
Alexandria, VA 22302-0268

Mr. Hei-Ki Dong  
Bell Communications Research  
Room PYA-1K207  
P.O. Box 1320  
Piscataway, NJ 08855-1320

Dr. Fritz Drasgow  
University of Illinois  
Department of Psychology  
603 E. Daniel St.  
Champaign, IL 61820

Defense Technical  
Information Center  
Cameron Station, Bldg. 5  
Alexandria, VA 22314  
(2 Copies)

Dr. Stephen Dunbar  
224B Lindquist Center  
for Measurement  
University of Iowa  
Iowa City, IA 52242

Dr. James A. Earles  
Air Force Human Resources Lab  
Brooks AFB, TX 78235

Dr. Susan Embretson  
University of Kansas  
Psychology Department  
426 Fraser  
Lawrence, KS 66045

Dr. George Englehard, Jr.  
Division of Educational Studies  
Emory University  
210 Fishburne Bldg.  
Atlanta, GA 30322

ERIC Facility Acquisitions  
2440 Research Blvd, Suite 550  
Rockville, MD 20850-3238

Dr. Benjamin A. Fairbank  
Operational Technologies Corp.  
5825 Callaghan, Suite 225  
San Antonio, TX 78228

Dr. Marshall J. Farr, Consultant  
Cognitive & Instructional Sciences  
2520 North Vernon Street  
Arlington, VA 22207

Dr. P.A. Federico  
Code 51  
NPRDC  
San Diego, CA 92152-6800

Dr. Leonard Feldt  
Lindquist Center  
for Measurement  
University of Iowa  
Iowa City, IA 52242

Dr. Richard L. Ferguson  
American College Testing  
P.O. Box 168  
Iowa City, IA 52243

Dr. Gerhard Fischer  
Leibgasse 5/3  
A 1010 Vienna  
AUSTRIA

Dr. Myron Fluchl  
U.S. Army Headquarters  
DAPE-MRR  
The Pentagon  
Washington, DC 20310-0300

Prof. Donald Fitzgerald  
University of New England  
Department of Psychology  
Armidale, New South Wales 2351  
AUSTRALIA

Mr. Paul Foley  
Navy Personnel R&D Center  
San Diego, CA 92152-6800

Dr. Alfred R. Freely  
APOSRL/NL, Bldg. 410  
Bolling AFB, DC 20332-6448

Dr. Robert D. Gibbons  
Illinois State Psychiatric Inst.  
Rm 529W  
1601 W. Taylor Street  
Chicago, IL 60612

Dr. Janice Gifford  
University of Massachusetts  
School of Education  
Amherst, MA 01003

Dr. Drew Gitomer  
Educational Testing Service  
Princeton, NJ 08541

Dr. Robert Glaeser  
Learning Research  
& Development Center  
University of Pittsburgh  
3939 O'Hara Street  
Pittsburgh, PA 15260

Dr. Sherrie Gott  
AFHRL/MOMJ  
Brooks AFB, TX 78235-5601

Dr. Bert Green  
Johns Hopkins University  
Department of Psychology  
Charles & 34th Street  
Baltimore, MD 21218

Michael Habon  
DORNIER GMBH  
P.O. Box 1420  
D-7990 Friedrichshafen 1  
WEST GERMANY

Prof. Edward Haertel  
School of Education  
Stanford University  
Stanford, CA 94305

Dr. Ronald K. Hambleton  
University of Massachusetts  
Laboratory of Psychometric  
and Evaluative Research  
Hills South, Room 152  
Amherst, MA 01003

Dr. Delwyn Harriach  
University of Illinois  
51 Gerty Drive  
Champaign, IL 61820

Dr. Grant Henning  
Senior Research Scientist  
Division of Measurement  
Research and Services  
Educational Testing Service  
Princeton, NJ 08541

Ms. Rebecca Hetter  
Navy Personnel R&D Center  
Code 63  
San Diego, CA 92152-6800

Dr. Thomas M. Hirsch  
ACT  
P. O. Box 168  
Iowa City, IA 52243

Dr. Paul W. Holland  
Educational Testing Service, 21-T  
Rosedale Road  
Princeton, NJ 08541

Dr. Paul Horst  
677 G Street, #184  
Chula Vista, CA 92010

Ms. Julia S. Hough  
Cambridge University Press  
40 West 20th Street  
New York, NY 10011

Dr. William Howell  
Chief Scientist  
AFHRL/CA  
Brooks AFB, TX 78235-5601

Dr. Lloyd Humphreys  
University of Illinois  
Department of Psychology  
603 East Daniel Street  
Champaign, IL 61820

Dr. Steven Hunka  
3-104 Educ. N.  
University of Alberta  
Edmonton, Alberta  
CANADA T6G 2G5

Dr. Huynh Huynh  
College of Education  
Univ. of South Carolina  
Columbia, SC 29208

Dr. Robert Jannarone  
Elec. and Computer Eng. Dept.  
University of South Carolina  
Columbia, SC 29208

Dr. Kumar Joag-dev  
University of Illinois  
Department of Statistics  
101 Illini Hall  
725 South Wright Street  
Champaign, IL 61820

Dr. Douglas H. Jones  
1280 Woodfern Court  
Toms River, NJ 08753

Dr. Brian Junker  
Carnegie-Mellon University  
Department of Statistics  
Schenley Park  
Pittsburgh, PA 15213

Dr. Michael Kaplan  
Office of Basic Research  
U.S. Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333-5600

Dr. Milton S. Katz  
European Science Coordination  
Office  
U.S. Army Research Institute  
Box 65  
FPO New York 09510-1500

Prof. John A. Keats  
Department of Psychology  
University of Newcastle  
N.S.W. 2308  
AUSTRALIA

Dr. Joo-keun Kim  
Department of Psychology  
Middle Tennessee State  
University  
P.O. Box 522  
Murfreesboro, TN 37132

Mr. Soon-Hoon Kim  
Computer-based Education  
Research Laboratory  
University of Illinois  
Urbana, IL 61801

Dr. G. Gage Kingsbury  
Portland Public Schools  
Research and Evaluation Department  
501 North Dixon Street  
P. O. Box 3107  
Portland, OR 97209-3107

Dr. William Koch  
Box 7246, Mesa, and Eval. Ctr.  
University of Texas-Austin  
Austin, TX 78703

Dr. Richard J. Koubek  
Department of Biomedical  
& Human Factors  
139 Engineering & Math Bldg.  
Wright State University  
Dayton, OH 45435

Dr. Leonard Kroeker  
Navy Personnel R&D Center  
Code 62  
San Diego, CA 92152-6800

Dr. Jerry Lehnus  
Defense Manpower Data Center  
Suite 400  
1600 Wilson Blvd  
Rosslyn, VA 22209

Dr. Thomas Leonard  
University of Wisconsin  
Department of Statistics  
1210 West Dayton Street  
Madison, WI 53705

Dr. Michael Levine  
Educational Psychology  
210 Education Bldg.  
University of Illinois  
Champaign, IL 61801

Dr. Charles Lewis  
Educational Testing Service  
Princeton, NJ 08541-0001

Mr. Rodney Lim  
University of Illinois  
Department of Psychology  
603 E. Daniel St.  
Champaign, IL 61820

Dr. Robert L. Linn  
Campus Box 249  
University of Colorado  
Boulder, CO 80309-0249

Dr. Robert Lockman  
Center for Naval Analysis  
4401 Ford Avenue  
P.O. Box 16268  
Alexandria, VA 22302-0268

Dr. Fredene M. Lord  
Educational Testing Service  
Princeton, NJ 08541

Dr. Richard Luechi  
ACT  
P. O. Box 168  
Iowa City, IA 52243

Dr. George B. Macready  
Department of Measurement  
Statistics & Evaluation  
College of Education  
University of Maryland  
College Park, MD 20742

Dr. Gary Marco  
Stop 31-E  
Educational Testing Service  
Princeton, NJ 08541

Dr. Clemen J. Martin  
Off - of Chief of Naval  
Operations (OP 13 F)  
Navy Annex, Room 2832  
Washington, DC 20350

Dr. James R. McBride  
HumRRO  
6430 Elmhurst Drive  
San Diego, CA 92120

Dr. Clarence C. McCormick  
HQ, USMEPCOM/MFPCT  
2500 Green Bay Road  
North Chicago, IL 60064

Mr. Christopher McCusker  
University of Illinois  
Department of Psychology  
603 E. Daniel St.  
Champaign, IL 61820

Dr. Robert McKinley  
Educational Testing Service  
Princeton, NJ 08541

Mr. Alan Mead  
c/o Dr. Michael Levine  
Educational Psychology  
210 Education Bldg.  
University of Illinois  
Champaign, IL 61801

Dr. Timothy Miller  
ACT  
P. O. Box 168  
Iowa City, IA 52243

Dr. Robert Mislevy  
Educational Testing Service  
Princeton, NJ 08541

Dr. William Montague  
NPRDC Code 13  
San Diego, CA 92152-6800

Ms. Kathleen Moreno  
Navy Personnel R&D Center  
Code 62  
San Diego, CA 92152-6800

Headquarters Marine Corps  
Code MPI-20  
Washington, DC 20380

Dr. Ratna Nandakumar  
Educational Studies  
Willard Hall, Room 213E  
University of Delaware  
Newark, DE 19716

Library, NPRDC  
Code P201L  
San Diego, CA 92152-6800

Librarian  
Naval Center for Applied Research  
in Artificial Intelligence  
Naval Research Laboratory  
Code 5510  
Washington, DC 20375-5000

Dr. Harold F. O'Neil, Jr.  
School of Education - WPH 801  
Department of Educational  
Psychology & Technology  
University of Southern California  
Los Angeles, CA 90089-0031

Dr. James B. Olsen  
WICAT Systems  
1875 South State Street  
Orem, UT 84058

Office of Naval Research,  
Code 1142CS  
800 N. Quincy Street  
Arlington, VA 22217-5000  
(6 Copies)

Dr. Judith Orasanu  
Basic Research Office  
Army Research Institute  
5001 Eisenhower Avenue  
Alexandria, VA 22333

Dr. Jesse Oransky  
Institute for Defense Analyses  
1801 N. Beauregard St.  
Alexandria, VA 22311

Dr. Peter J. Pashley  
Educational Testing Service  
Rosedale Road  
Princeton, NJ 08541

Wayne M. Patience  
American Council on Education  
GED Testing Service, Suite 20  
One Dupont Circle, NW  
Washington, DC 20036

Dr. James Paulson  
Department of Psychology  
Portland State University  
P.O. Box 751  
Portland, OR 97207

Dept. of Administrative Sciences  
Code 54  
Naval Postgraduate School  
Monterey, CA 94043-5026

Dr. Mark D. Reckase  
ACT  
P. O. Box 168  
Iowa City, IA 52243

Dr. Malcolm Ree  
AFHRL/MOA  
Brooks AFB, TX 78235

Mr. Steve Reas  
N660 Elliott Hall  
University of Minnesota  
75 E. River Road  
Minneapolis, MN 55455-0344

Dr. Carl Rom  
CNET-PDCID  
Building 90  
Great Lakes NTC, IL 60088

Dr. J. Ryan  
Department of Education  
University of South Carolina  
Columbia, SC 29208

Dr. Fumiko Samejima  
Department of Psychology  
University of Tennessee  
310B Austin Peay Bldg.  
Knoxville, TN 37916-0400

Mr. Drew Sands  
NPRDC Code 62  
San Diego, CA 92152-6800

Lowell Schoer  
Psychological & Quantitative  
Foundations  
College of Education  
University of Iowa  
Iowa City, IA 52242

Dr. Mary Schmitt  
4100 Parkside  
Carlsbad, CA 92008

Dr. Don Sepall  
Navy Personnel R&D Center  
San Diego, CA 92152

Dr. Robin Sheely  
University of Illinois  
Department of Statistics  
101 Illini Hall  
725 South Wright St.  
Champaign, IL 61820

Dr. Kazuo Shigematsu  
7-9-24 Kugenuma-Kaigan  
Fujisawa 251  
JAPAN

Dr. Randall Shumaker  
Naval Research Laboratory  
Code 5510  
4555 Overlook Avenue, S.W.  
Washington, DC 20375-5000

Dr. Richard E. Snow  
School of Education  
Stanford University  
Stanford, CA 94305

Dr. Richard C. Sorenson  
Navy Personnel R&D Center  
San Diego, CA 92152-6800

Dr. Judy Spray  
ACT  
P.O. Box 168  
Iowa City, IA 52243

Dr. Martha Stocking  
Educational Testing Service  
Princeton, NJ 08541

Dr. Peter Stokoff  
Center for Naval Analysis  
4401 Ford Avenue  
P.O. Box 16268  
Alexandria, VA 22302-0268

Dr. William Stout  
University of Illinois  
Department of Statistics  
101 Illini Hall  
725 South Wright St.  
Champaign, IL 61820

Dr. Hanbaran Swaminathan  
Laboratory of Psychometric and  
Evaluation Research  
School of Education  
University of Massachusetts  
Amherst, MA 01003

Mr. Brad Symptom  
Navy Personnel R&D Center  
Code 62  
San Diego, CA 92152-6800

Dr. John Tangney  
AFOSR/NL, Bldg. 410  
Bolling AFB, DC 20332-6448

Dr. Kikumi Tatsuoka  
Educational Testing Service  
Mail Stop 03-T  
Princeton, NJ 08541

Dr. Maurice Tatsuka  
Educational Testing Service  
Mail Stop 03-T  
Princeton, NJ 08541

Dr. David Thissen  
Department of Psychology  
University of Kansas  
Lawrence, KS 66044

Mr. Thomas J. Thomas  
Johns Hopkins University  
Department of Psychology  
Charles & 34th Street  
Baltimore, MD 21218

Mr. Gary Thomason  
University of Illinois  
Educational Psychology  
Champaign, IL 61820

Dr. Robert Tsutakawa  
University of Missouri  
Department of Statistics  
222 Math. Sciences Bldg.  
Columbia, MO 65211

Dr. Ledyard Tucker  
University of Illinois  
Department of Psychology  
603 E. Daniel Street  
Champaign, IL 61820

Dr. David Vale  
Assessment Systems Corp.  
2233 University Avenue  
Suite 440  
St. Paul, MN 55114

Dr. Frank L. Vicino  
Navy Personnel R&D Center  
San Diego, CA 92152-6800

Dr. Howard Wainer  
Educational Testing Service  
Princeton, NJ 08541

Dr. Michael T. Weller  
University of Wisconsin-Milwaukee  
Educational Psychology Department  
Box 413  
Milwaukee, WI 53201

Dr. Ming-Mei Wang  
Educational Testing Service  
Mail Stop 03-T  
Princeton, NJ 08541

Dr. Thomas A. Warm  
FAA Academy AAC934D  
P.O. Box 25082  
Oklahoma City, OK 73125

Dr. Brian Waters  
HumRRO  
1100 S. Washington  
Alexandria, VA 22314

Dr. David J. Weiss  
Néé Elliott Hall  
University of Minnesota  
75 E. River Road  
Minneapolis, MN 55455-0344

Dr. Ronald A. Weitzman  
Box 146  
Carmel, CA 93921

Major John Welsh  
AFHRL/MOAN  
Brooks AFB, TX 78223

Dr. Douglas Wetzel  
Code 51  
Navy Personnel R&D Center  
San Diego, CA 92152-6800

Dr. Rand R. Wilson  
University of Southern  
California  
Department of Psychology  
Los Angeles, CA 90089-1061

German Military Representative  
ATTN: Wolfgang Widgrub  
Streitkräfteamt  
D-5308 Bonn 2  
4000 Brandywine Street, NW  
Washington, DC 20016

Dr. Bruce Williams  
Department of Educational  
Psychology  
University of Illinois  
Urbana, IL 61801

Dr. Hilda Wing  
Federal Aviation Administration  
800 Independence Ave. SW  
Washington, DC 20591

Mr. John H. Wolfe  
Navy Personnel R&D Center  
San Diego, CA 92152-6800

Dr. George Wong  
Biostatistics Laboratory  
Memorial Sloan-Kettering  
Cancer Center  
1275 York Avenue  
New York, NY 10021

Dr. Wallace Wulfeck, III  
Navy Personnel R&D Center  
Code 51  
San Diego, CA 92152-6800

Dr. Kentaro Yamamoto  
02-T  
Educational Testing Service  
Rosedale Road  
Princeton, NJ 08541

Dr. Wendy Yen  
CTB/McGraw Hill  
Del Monte Research Park  
Monterey, CA 93940

Dr. Joseph L. Young  
National Science Foundation  
Room 320  
1800 G Street, N.W.  
Washington, DC 20550

Mr. Anthony R. Zarr  
National Council of State  
Boards of Nursing, Inc.  
625 North Michigan Avenue  
Suite 1544  
Chicago, IL 60611